DOCUMENT RESUME

ED 472 876 IR 058 581

AUTHOR	Calvert, Philip J.; Gorman, G.E.
TITLE	Testing Article Quality in LIS Journals: The Search Continues.
PUB DATE	2002-08-00
NOTE	9p.; In: Libraries for Life: Democracy, Diversity, Delivery. IFLA Council and General Conference: Conference Programme and Proceedings (68th, Glasgow, Scotland, August 18-24, 2002); see IR 058 549.
AVAILABLE FROM	For full text: http://www.ifla.org.
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Evaluation Criteria; Evaluation Methods; Foreign Countries; *Information Science; *Journal Articles; *Library Science; Periodicals; Pilot Projects; *Scholarly Journals
IDENTIFIERS	*Library Journals

ABSTRACT

During the first half of 2000 a pilot project on library and information science (LIS) journal quality was undertaken in the Asian region for the International Federation of Library Associations and Institutions (IFLA) Round Table on Library and Information Science Journals (RTLISJ). It showed that a simple, but effective, methodology could be used to determine the most commonly accepted criteria for evaluating LIS journal quality, the success or otherwise of LIS journals in meeting those criteria, and some critical factors that could improve the quality of those journals. The current project, developed on the basis of the pilot study, is examining a selective random sample of LIS journals from around the world to determine their quality using the criteria devised and tested in the pilot study. This paper reports on the findings to date, which,are currently being completed and remain to be compared with the views of editors and editorial board members on the quality of their respective journals. Preliminary findings suggest that independent qualitative assessment of journal content does not match the perceptions of those intimately involved with controlling this content. Includes five tables. (Contains 14 references.) (Author)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 472 876

68th IFLA Council and General Conference
August 18-24, 2002

Code Number: 169-118-E
Division Number: VII
Professional Group: Library and Information Science Journals
Joint Meeting with: -
Meeting Number: 118
Simultaneous Interpretation: -

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

S. Koopman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Testing Article Quality in LIS Journals: The Search Continues

Philip J. Calvert
and

G.E. Gorman
School of Information Management
Victoria University of Wellington

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Abstract

During the first half of 2000 a pilot project on library and information science (LIS) journal quality was undertaken in the Asian region for the IFLA Round Table on Library and Information Science Journals (RTLISJ). It showed that a simple but effective methodology could be used to determine the most commonly acepted criteria for evaluating LIS journal quality, the success or otherwise of LIS journals in meeting those criteria, and some critical factors that could improve the quality of those journals. The present project, developed on the basis of the pilot study, is examining a selective random sample of LIS journals from around the world to determine their quality using the criteria devised and tested in the pilot study. This paper reports on the findings to date, which are currently being completed and remain to be compared with the views of editors and editorial board members on the quality of their respective journals. Preliminary findings suggest that independent qualitative assessment of journal content does not match the perceptions of those intimately involved with controlling this content.

Introduction

We are all familiar with the traditional approach to measuring journal 'quality' against quantitative measures such as circulation, total number of pages per volume, number of times cited in the literature, coverage by indexing services, etc. Such approaches have been firmly embedded in our profession since the important work by Eugene Garfield on citation analysis, and continues almost unquestioned down to

IR058581

BEST COPY AVAILABLE

the present day, as for example in the recent paper by Ronald Rousseau, which states in part, 'perhaps this overview will inspire fellow scientists to construct an overall model explaining observed citation scores, and hence lead to a better understanding of their role in institutional and national evaluations'. (Rousseau 2002) This is despite other work which questions the real value of quantitative measures in determining the quality and usefulness of journal literature, as raised in some of the studies by Altmann and Gorman (1999).

This investigation is based on the premise that, while these measures are perhaps appropriate for determining 'leading' journals or for ranking journals within a discipline, they are clearly lacking when it comes to determining the intrinsic quality of individual journals. It adopts a more qualitative approach to understanding journal quality based on perceptions of key stakeholders – in particular readers, editors and editorial board members. The intention is to help establish standards of excellence among LIS journals, so that they can become more effective channels for the communication of theory and practice to the various information professions within LIS.

The pilot research project in 2000, undertaken on behalf of IFLA's Round Table of Library and Information Science Journals (RTLISJ), followed a presentation by one of the investigators at IFLA Bangkok which sought to establish general criteria for assessing journal submissions from Asian authors. (Gorman 1999) That paper suggested that quantitative measures were perhaps unsuitable for evaluating the qualitative factors that contribute to journal excellence. Following that presentation, the pilot study of journal quality took a more qualitative approach to understanding journal quality based on perceptions of key stakeholders – the editors, referees, editorial board members and impartial readers. The specific target of the study was Library and Information Science (LIS) journals published in Asia.

Essentially, this study was meant to be a pilot project for a more broadly based international investigation of journal quality. There is a widespread view that LIS journals around the world are of somewhat uneven quality, that they do not meet a common set of standards of excellence, that there may not even be such a set of standards, that journals in developing countries in particular might benefit from better understanding of quality, and that all such journals can become more effective channels for the communication of theory and practice to the various information professions within LIS.

'Improved journal quality' must take into account the reality that contributors to journals write for a variety of reasons (Gorman 1999). Most commonly, we write to disseminate new research findings or ideas. The publication of a paper establishes precedents in the formation of new knowledge, and it puts the new information in the professional domain where it can be scrutinised, criticised and either accepted or rejected. It may then contribute to further discourse. The author also makes personal gains by adding to a list of publications that can be used for tenure and promotion, for gaining professional acceptance that may lead to speaking engagements, consultancy work, perhaps even awards.

There is, then, an apparent contradiction between the intrinsic and extrinsic reasons for scholarly and professional writing, for personal benefits can result from numerous papers of indifferent quality, whereas the imperative for the discipline is the discovery of new conceptual approaches and new techniques, for which the need is for papers of the highest quality. That the scholarly communication system has survived almost unchanged for so long shows its robustness, but the inherent contradictions in the system make it vulnerable to distortion under certain circumstances. If, for example, the 'publish or perish' imperative creates such demand among hopeful authors that editors are overwhelmed with manuscripts of an indifferent quality, then there is potential for the erosion of standards. This may occur if new journals start up to cater for the unfulfilled demand from hopeful authors. There is some emerging evidence, from current research by Calvert and Shi (2000) on quality and quantity in journal publishing, that this has happened in China already.

Both extrinsic and intrinsic reasons lead to publications that might be assessed by quantitative means, but also by qualitative means. The fact that paper *x* is cited *y* times is not an indication of quality, but rather that it is cited – it is available, it is in a journal held by many libraries, the author (or publisher or editor) is particularly good at self-promotion. Behind the quantifiable factors, then, are as yet untested qualitative factors, which is what has led to the present project.

Methodology

From our study of the literature on journal quality, especially LIS journal quality (the literature is reviewed in Gorman and Calvert 2001), several criteria were selected as most relevant and applicable without the need for extensive explanation as to meaning, and these criteria were seen to fall into three major categories:

- prestige (of the editor, etc.)
- properties of articles within a journal
- presentational aspects.

The present project has focussed on the second of the larger categories (properties of articles), because we felt this was likely to be the most crucial in determining journal quality. Accordingly, only the disaggregated properties of journal articles are being examined in detail.

Gorman (1999) had suggested six criteria for the evaluation of submissions to Asian LIS journals, and it was these six that the editors agreed ultimately accepted as most relevant to the assessment of paper quality/content:

- advancement of knowledge
- new information or data
- theoretical soundness
- level of scholarship
- acceptable research design
- appropriate methodology and analysis.

These six criteria were used for the next stage of the project, which involved in-depth analysis of the sample copies supplied by the editors. Papers were chosen at random for examination, five papers from each journal being the norm. Once a paper had been read, a 'score' was given (from 0 to 10, with 10 being the highest standard) based on how well it had measured against each specific criterion. The purpose was not to assess the journals, but rather to compare the criteria, so an average score was calculated for each of the six criteria, and then they were placed in a ranked list. The two highest rated categories (new information or data, acceptable research design) scored well ahead of the others. The third category (level of scholarship) was on its own in third place, with the other three grouped together at the bottom. From this analysis, the criteria can be placed in descending order of importance as follows:

- new information or data
- acceptable research design
- level of scholarship
- theoretical soundness
- advancement of knowledge
- appropriate methodology and analysis.

Table 1. Regional Distribution of Journals Being Analysed, Aug-Dec 2002

Region	Number
North America	60
Latin America	5
Europe	85
Africa and the Middle East	10
Pacific Islands	5
Australia and New Zealand	5
Asia	30
Total	***200***

Findings to Date

As noted in the introductory discussion, quantitative evaluation of journal quality has usually focussed on citation analysis. By contrast, this research examines the disaggregated properties of journal articles using qualitative methods as outlined in the discussion of methodology.

To date the research team has examined LIS journals from 10 countries listed in Table 2. It will be seen that all of these are Western countries, because acquiring sample issues from non-Western countries is taking longer than anticipated. However, we are assured by the editors that many dozens of these journals are currently on their way to us.

Table 2. Number of Journals Examined, August 2002

Country	*Number of Journals*
USA	*14*
UK	*14*
Canada	*3*
Australia	*3*
Denmark	*1*
France	*2*
Netherlands	*1*
New Zealand	*1*
Sweden	*1*
Total	40

The researchers have read at least one article from each journal. These are always the first lengthy paper in the hard copy, or the first listed contribution in the content of an electronic journal. Each article receives a score on a scale of 1-10 for each of the four criteria identified by editors as key factors in the pilot study: currency of information, research design (where applicable), scholarly quality, and soundness of theory.

Is this approach subjective? Yes, indeed it is, just as readers make their subjective judgements, and so do editors and referees. Remember that in our view quality is not a matter of citation counts but of personal interaction with the content of a paper. Therefore, the research team make no apology for the subjectiveness of the marking. Scholarly quality, for example, has to be a subjective judgement, yet it is this very characteristic that editors say they are looking for above all others. What we look for specifically in terms of scholarly quality are two features:

quality of analysis applied to the content
the author's ability to generate new knowledge.

For theoretical soundness we wanted to see some evidence of the use of theory, at the very least, and at best the use of multiple theories appropriate to the context.

Word Count and Article Quality

Having given each article a score, what could this be compared to that enlightened us about journal quality? Having some experience of journal editing, we made an intuitive choice of one other measure to use in the analysis, that of word count. This made the assumption that the longer the article, the more it would exhibit some aspect of journal quality or another. Bearing in mind the small sample size to date, this does seem to be shown by the data so far.

The articles have been divided into three groups:

<2000 words
2000-4000 words
>4000 words.

The first group was slightly smaller than the other two, and the other two groups were equal in size.

Table 3. Size as Related to Scholarly Quality

Size	Quality Factor
<2000 words	3.4
2000-4000 words	7.8
>4000 words	7.3

The scores for scholarly quality produced an indication (and we emphasise that this is a tentative indication at present) that scholarly quality is more likely to be found in article over 2000 words long. As Table 3 shows, the average score for papers under 2000 words was 3.4, whereas it was 7.8 for the middle group and 7.3 if the paper was longer than 4000 words. Most of the short articles were reports on new developments, the 'how-we-run-our-library-good' type of article, or simply opinion pieces. The quality of thinking was often sadly lacking, and it makes us wonder how much new developments are analysed before they are reported in the LIS literature. In contrast, the longer articles were often the product of clear, reflective thinking by their authors, and this in our view is the kind of paper that adds quality to a journal.

Table 4. Size as Related to Theoretical Soundness

Size	Theory Factor
<2000 words	1.5
2000-4000 words	4.8
>4000 words	6.6

Something similar was found with the scores for theoretical soundness, though in this case the scores rose with each step up the word length ladder. The short articles were almost completely devoid of theory (they scored only 1.5 on average), the middle length articles scored 4.8, which is still rather too low, but the longer papers averaged 6.6. Overall these scores were disappointing, and though it seems likely that

longer articles will include some relevant theory, the use of theory is clearly not a strong point in the LIS literature. We must then ask why editors rank this as a key criterion, when the submissions they approve for publication are sorely lacking in this regard.

Table 5. Size as Related to Currency of Information

Size	Currency Factor
<2000 words	7.0
2000-4000 words	6.3
>4000 words	6.8

By contrast, the scores for currency of information showed little difference no matter the length of the article. The shortest papers scored the highest on this criterion (7.0) compared with 6.3 for the middle group and 6.8 for the longest papers. This suggests that editors seeking current information should encourage short articles, presumably because they can be written quickly! Longer articles incorporating theory and critical thinking will take longer to write and so are less likely to be completely current.

The scores for research design showed little difference, but it needs to be said that there were a disappointing number of research papers in the total sample, and only one in a paper shorter than 2000 words. There was no real difference in scores for research method in the two other groups.

Word Count and Peer Review

Three simple factors were then used to divide the articles and the simple measure of word count used as a means of assessing if the patterns were different.

The first factor used was whether the article was peer reviewed or not. This is commonly used in journal evaluation (some refs). Though the sample size so far is not large enough to warrant detailed statistical analysis, very simple descriptive statistics suggest that there is a difference based solely on word length. Refereed papers averaged about 3600 words, while the non-refereed papers averaged only 2500 words. It should be noted that some very long papers raised the average in both cases, and the median word length for refereed papers was 3250 words and for non-refereed papers the median length was 2000 words. While being aware of the word 'significant', it seems quite clear that there is a difference between refereed and non-refereed papers.

Word Count and Type of Article

The second factor assessed was whether the article was either a report of research or a practical paper that was mainly descriptive in content. There were a few borderline cases that needed a judgement call on this. The practical papers averaged 2600 words while the research papers averaged 3600 words showing that once again there was a measurable difference between the two. The median of practical papers was 2300 words and the median of research papers was 3000 words.

Word Count and Type of Publisher

The third factor used was whether a commercial entity or a non-commercial one published the journal containing the article. In most cases the non-commercial publishers are associations of librarians, though publications of national libraries and similar institutions were included in this category because their primary purpose is not-for-profit. The results showed no real difference between the two. Articles in commercial journals averaged just over 3000 words, which was the same as for non-commercial journals.

BEST COPY AVAILABLE

The median length for commercial journals was 3000 words, and for non-commercial journals it was only slightly less.

Summary to Date

Summing up the results so far, there appears to be some justification for further experimentation with word length as a unit of measure in journal quality. Longer articles show higher scores on the two important criteria of scholarly quality and theoretical soundness. On the other hand, the LIS literature is very full of short papers exhibiting no great evidence of critical thinking or the use of relevant theory.

Having established the standards, it is then proposed to offer an IFLA workshop on journal quality, especially for editors of journals in developing countries, enabling them to share experiences with experienced editors from elsewhere.

References

Ali, S.N.; Young, H.C.; and Ali, N.M. (1996) 'Determining the Quality of Publications and Research for Tenure or Promotion Decisions: A Preliminary Checklist to Assist'. *Library Review* 45, 1: 39 – 53.

Altmann, K.G., and Gorman, G.E. (1999) 'The Relevance of "Cited by Leading Journal" to Serials Management in Australian University Libraries'. *Australian Library Journal* 48, 2: 101-115

Anderson, P. (1997) '"Gatekeepers" and the Quality of the Journal Literature: Findings from a Survey of Journal Editors into the Issue of Alleged Excessive Publication in Scholarly and Scientific Journals'. *Serials Review* 23, 2: 45-57.

Calvert, P.J., and Shi, Z. Quality versus Quantity: Contradictions in LIS Journal Publishing in China'. *New Library World* 2000 (forthcoming).

Chressanthis, G.A., and Chressanthis, J.D. (1993) 'The Relationship between Manuscript Submission Fees and Journal Quality'. *Serials Librarian* 24, 1: 71-86.

Day, A., and Peter, J. (1994) 'Quality Indicators in Academic Publishing'. *Library Review* 43, 7: 4 – 72.

Garfield, E. (1977) 'Citation Frequency as a Measure of Research Activity'. In *Essays of an Information Scientist. Volume 1, 1962-1973*, 406-408. Philadelphia: ISI Press.

Gorman, G.E (1999) 'Library and Information Science Journals in the Asian Context'. *65th IFLA Council and General Conference, August 20-28, 1999: Papers and Proceedings.* The Hague: IFLA, 1999. Code 005-118-E. Booklet VII.

Gorman, G.E., and Calvert, P.J. (2000) 'Journal Quality in the Asian Region: Results of a Pilot Study for the IFLA Round Table of Library and Information Science Journals'. *66th IFLA Council and General Conference, Jerusalem, Israel, 13-18 August 2000: Conference Proceedings.* The Hague: IFLA, 2000 http://www.ifla.org.sg/iv/ifla66/papers/156-125e.htm. Also published as Gorman, G.E., and Calvert, P.J. (2001) 'Journal Quality in the Asian Region: Results of a Pilot Study for the IFLA Round Table of Library and Information Science Journals'. *The Serials Librarian* 41, 1: 99-112.

Lester, J.P. (1990) 'Evaluating the Evaluators: Accrediting Knowledge and the Ranking of Political Science Journals'. *PS: Political Science and Politics* 23: 445-447.

Nisonger, T.E. (1996) 'Authorship in *Library Acquisitions: Practice and Theory*'. *Library Acquisitions: Practice and Theory* 20: 395-418.

Nkereuwem, E.E. (1997) 'Accrediting Knowledge: The Ranking of Library and Information Science Journals'. *Library Review* 46, 2: 99 - 104.

Rousseau, R. (2002) 'Journal Evaluation: Technical and Practical Issues'. *Library Trends* 50, 3: 418-439.

Sen, B.K. (1999) 'Changes in Impact Factor'. *Malaysian Journal of Library and Information Science* 4, 2: 41-46.

About the Authors

The principal investigators, Dr G E Gorman and Mr Philip Calvert, are Professor and Senior Lecturer respectively in the School of Information Management at Victoria University of Wellington and experienced editors. Calvert is former Editor of *New Zealand Libraries* and currently reviews editor of *Online Information Review* and *The Electronic Library*. Gorman, Secretary of the IFLA RSCAO and Chair of the IFLA Section on Library and Information Science Journals, is Editor of *Online Information Review*, Associate Editor of *Library Collections, Acquisitions and Technical Services*, Editor of *Library LINK* and Regional Editor (Asia-Pacific) Library and Information Management for Emerald/MCB. E-mail: gary.gorman@vuw.ac.nz or philip.calvert@vuw.ac.nz.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

X	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").